

Appendix I: Alignment of instant SEQ ID NO: 1 and Q9H4I2, publicly available June 2002

BLASTP 2.2.22+

Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Reference for compositional score matrix adjustment: Stephen F. Altschul, John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schaffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", FEBS J. 272:5101-5109.

RID: KMB3KFZ911R

Query= SID_1
Length=956

Sequences producing significant alignments:	Score (Bits)	E Value
lcl 22739 TrEMBL Q9H4I2 Release 21 01-JUN-2002	1989	0.0

ALIGNMENTS

>lcl|22739 TrEMBL|Q9H4I2|Release 21|01-JUN-2002
Length=956

Score = 1989 bits (5152), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 956/956 (100%), Positives = 956/956 (100%), Gaps = 0/956 (0%)

Query	1	MASKRKSTTPCMIPVKTVVLQDASMEAQPAETLPEGPQQDLPEASAASSEAAQNPSSTD	60
		MASKRKSTTPCMIPVKTVVLQDASMEAQPAETLPEGPQQDLPEASAASSEAAQNPSSTD	
Sbjct	1	MASKRKSTTPCMIPVKTVVLQDASMEAQPAETLPEGPQQDLPEASAASSEAAQNPSSTD	60
Query	61	GSTLANGHRSTLDGYLYSCKYCDFRSHDMTQFVGHMNSEHTDFNKDPTFVCSGCSFLAKT	120
		GSTLANGHRSTLDGYLYSCKYCDFRSHDMTQFVGHMNSEHTDFNKDPTFVCSGCSFLAKT	
Sbjct	61	GSTLANGHRSTLDGYLYSCKYCDFRSHDMTQFVGHMNSEHTDFNKDPTFVCSGCSFLAKT	120
Query	121	PEGLSLHNATCHSGEASFVWNVAKPDNHVVVEQSIPESTSTPDLAGEPSAEGADGQAEII	180
		PEGLSLHNATCHSGEASFVWNVAKPDNHVVVEQSIPESTSTPDLAGEPSAEGADGQAEII	
Sbjct	121	PEGLSLHNATCHSGEASFVWNVAKPDNHVVVEQSIPESTSTPDLAGEPSAEGADGQAEII	180
Query	181	ITKTPIMKIMKGKAEAKKIHTLKENVPSQPVGEALPKLSTGEMEVREGDHSFINGAVPVS	240
		ITKTPIMKIMKGKAEAKKIHTLKENVPSQPVGEALPKLSTGEMEVREGDHSFINGAVPVS	
Sbjct	181	ITKTPIMKIMKGKAEAKKIHTLKENVPSQPVGEALPKLSTGEMEVREGDHSFINGAVPVS	240
Query	241	QASASSAKNPHAANGPLIGTVPVLPAGIAQFLSLQQQPPVHAQHVVHQPLPTAKALPKVM	300
		QASASSAKNPHAANGPLIGTVPVLPAGIAQFLSLQQQPPVHAQHVVHQPLPTAKALPKVM	
Sbjct	241	QASASSAKNPHAANGPLIGTVPVLPAGIAQFLSLQQQPPVHAQHVVHQPLPTAKALPKVM	300
Query	301	IPLSSIPTYNAAMDSNSFLKNSFHKFPYPTKAELCYLTVVTKYPEEQWKIWFQAQRLKQG	360
		IPLSSIPTYNAAMDSNSFLKNSFHKFPYPTKAELCYLTVVTKYPEEQWKIWFQAQRLKQG	
Sbjct	301	IPLSSIPTYNAAMDSNSFLKNSFHKFPYPTKAELCYLTVVTKYPEEQWKIWFQAQRLKQG	360
Query	361	ISWSPEEIEDARKKMFNTVIQSVQPTITVLNTPLVASAGNVQHLLQAALPGHVVGQPEG	420

Sbjct	361	ISWSPEEIEDARKKMFNTVIQSVQPQTITVLNTPPLVASAGNVQHLLQAALPGHVVGQPEG ISWSPEEIEDARKKMFNTVIQSVQPQTITVLNTPPLVASAGNVQHLLQAALPGHVVGQPEG	420
Query	421	TGGGLLVLTQPLMANGQLATSSPLPLTVTSVPKQPGVAPINTVCSNTTSAVKVVNAQSL TGGGLLVLTQPLMANGQLATSSPLPLTVTSVPKQPGVAPINTVCSNTTSAVKVVNAQSL	480
Sbjct	421	TGGGLLVLTQPLMANGQLATSSPLPLTVTSVPKQPGVAPINTVCSNTTSAVKVVNAQSL	480
Query	481	TACPSITSQAFLDASIYKNKKSHEQLSALKGSFCRNQFPQGSEVEHLTKVTGLSTREVRK TACPSITSQAFLDASIYKNKKSHEQLSALKGSFCRNQFPQGSEVEHLTKVTGLSTREVRK	540
Sbjct	481	TACPSITSQAFLDASIYKNKKSHEQLSALKGSFCRNQFPQGSEVEHLTKVTGLSTREVRK	540
Query	541	WFSDDRYHCRNLKGSRAMIPGDHSSIIIDSVPVVSFSPSSKVPEVTCIPTTATLATHPSA WFSDDRYHCRNLKGSRAMIPGDHSSIIIDSVPVVSFSPSSKVPEVTCIPTTATLATHPSA	600
Sbjct	541	WFSDDRYHCRNLKGSRAMIPGDHSSIIIDSVPVVSFSPSSKVPEVTCIPTTATLATHPSA	600
Query	601	KRQSWHQTPDFTPTKYKERAPEQLRALESSFAQNPLPLDEELDRLRSETKMTRREIDSWF KRQSWHQTPDFTPTKYKERAPEQLRALESSFAQNPLPLDEELDRLRSETKMTRREIDSWF	660
Sbjct	601	KRQSWHQTPDFTPTKYKERAPEQLRALESSFAQNPLPLDEELDRLRSETKMTRREIDSWF	660
Query	661	SERRKKVNAEETKKAENASQEEEEAAEDEGGEEDLASELRVSGENGSLEMPSSHILAER SERRKKVNAEETKKAENASQEEEEAAEDEGGEEDLASELRVSGENGSLEMPSSHILAER	720
Sbjct	661	SERRKKVNAEETKKAENASQEEEEAAEDEGGEEDLASELRVSGENGSLEMPSSHILAER	720
Query	721	KVSPKINLKNLRVTEANGRNEIPGLGACDPEDDESINKLAELPGKVSCKKTAQQRHLLR KVSPKINLKNLRVTEANGRNEIPGLGACDPEDDESINKLAELPGKVSCKKTAQQRHLLR	780
Sbjct	721	KVSPKINLKNLRVTEANGRNEIPGLGACDPEDDESINKLAELPGKVSCKKTAQQRHLLR	780
Query	781	QLFVQTQWPSNQDYDSIMAQTGLPRPEVVRWFGDSRYALKNGQLKWEYEDYKRGNFPPGLL QLFVQTQWPSNQDYDSIMAQTGLPRPEVVRWFGDSRYALKNGQLKWEYEDYKRGNFPPGLL	840
Sbjct	781	QLFVQTQWPSNQDYDSIMAQTGLPRPEVVRWFGDSRYALKNGQLKWEYEDYKRGNFPPGLL	840
Query	841	VIAPGNRELLQDYIMTHKMLYEEDLQNLCDKTQMSSQVQKQWFAEKMGEETRAVADTGSE VIAPGNRELLQDYIMTHKMLYEEDLQNLCDKTQMSSQVQKQWFAEKMGEETRAVADTGSE	900
Sbjct	841	VIAPGNRELLQDYIMTHKMLYEEDLQNLCDKTQMSSQVQKQWFAEKMGEETRAVADTGSE	900
Query	901	DQGPGTGELTAVHKMGMDTYSEVSENSESWEPRVPEASSEPFDTSSPQAGRQLETD DQGPGTGELTAVHKMGMDTYSEVSENSESWEPRVPEASSEPFDTSSPQAGRQLETD	956
Sbjct	901	DQGPGTGELTAVHKMGMDTYSEVSENSESWEPRVPEASSEPFDTSSPQAGRQLETD	956